

---

## Rule DAS400: Access characteristics of significant data sets

---

**Finding:** CPExpert identifies the access characteristics of data sets managed by DFSMS (and reported in SMF Type 42 records), for the volumes that have the most potential for performance improvement.

**Impact:** This finding is used to assess the data sets that reside on the volumes with significant performance improvement potential.

**Logic flow:** This is a basic finding. There are no predecessor rules.

**Discussion:** If SMF Type 42 (Data Set Statistics) information is available<sup>1</sup> in a MXG performance data base, the DASD Component will process the MXG TYPE42DS file. The DASD Component select information describing data sets that were referenced during RMF measurement intervals in which the poorly performing device(s) exceeded the average performance. The result from this analysis is displayed in Rule DAS400.

There are several considerations with respect to the information produced by Rule DAS400:

- There can be many data sets referenced on the devices (hundreds or even thousands of data sets can be referenced). It is not helpful to have a large number of data sets listed. Consequently, CPExpert provides two guidance variables to control how many data sets are listed by Rule DAS400:
  - The **LIST42DS** variable can be used to limit the number of data sets listed in any RMF measurement interval, for a particular device. Only the number of data sets specified by the LIST42DS variable will be listed individually, and information about any remaining data sets will be summarized and listed on a single line.

The default value for the LIST42DS variable is 10, indicating that 10 data sets will be listed for each RMF interval in which a poorly-performing device had a performance problem. You can alter this number of data sets listed by using the LIST42DS guidance variable.

- Some data sets might have a low I/O activity and would not be interesting to analyze. CPExpert provides the **MIN42PCT** variable to limit the number of data sets listed in any RMF measurement interval,

---

<sup>1</sup>%LET TYPE42DS = Y; was specified in USOURCE(GENGUIDE).

---

for a particular device. Only data sets having a percent of activity for the total volume greater than the percent specified by the MIN42PCT variable will be listed individually, and any remaining data sets will be summarized and listed on a single line.

The default value for the MIN42PCT variable is 0.1, indicating that data sets will not be listed individually for any RMF interval unless the data set intensity of access (I/O rate \* response time) was greater than 0.1% of the total volume intensity of access. You can alter this percent of data sets listed by using the MIN42PCT guidance variable.

For example, if you specified %LET MIN42PCT = 25, a maximum of 4 data sets would be listed in any RMF interval since no more than 4 data sets could have 25% or higher access intensity.

Please note that, regardless of the data set access intensity of any particular data set, only the number of data sets specified by the **LIST42DS** guidance variable (described earlier) will be listed. This means that there are two ways to limit the number of data sets listed: (1) the LIST42DS which limits the number of data sets listed in any RMF interval, and (2) the **MIN42PCT** guidance variable which limits the data sets listed to those that exceed the specified percent.

- Some devices might have data sets that are not managed by DFSMS. In this case, information would not be available in SMF Type 42 records and the data sets could not be reported by CPExpert.
- Depending on the timing in which SMF writes Type 42 records, some information might not be represented in the RMF Type 74 intervals by which the DASD Component analyzes DASD performance problems. Consequently, the information presented in Rule DAS400 might not correspond to the information presented in other rules (for example, might not correspond to information presented in Rule DAS100).
- SMF writes Type 42 (Data Set Statistics) records when (1) a DASD data set is closed, or (2) immediately after the recording of the SMF Type 30 interval record. There are two implications of this SMF write characteristic:
  - A data set likely would be closed at some time between SMF writing Type 30 interval records. Consequently, some Type 42 (Data Set Statistics) records would not correspond to the SMF Type 30 interval, and the data set information reported by the DASD Component might not correspond to the information reported for the device by other rules.
  - The SMF Type 30 recording interval controls when Type 42 (Data Set Statistics) records are written (for those data sets that are OPEN when

---

the Type 30 record is written). However, the SMF Type 30 recording interval might not correspond with the RMF recording interval. If the SMF and RMF writing is synchronized, the SMF Type 30 recording interval would be synchronized with the RMF recording interval. However, the synchronization option might not have been selected.

Even if the synchronization option had been selected, a different recording interval frequency could have been selected for Type 30 and RMF recording. For example, RMF recording could occur at 15 minute intervals, but SMF Type 30 recording could occur at 30 minute intervals. While the intervals could be synchronized, there would be twice as many RMF recording intervals as there were SMF Type 30 (and Type 42) recording intervals.

The SMF and RMF recording considerations might cause consternation if a user were not aware of the potential problems. If the SMF and RMF recording intervals are synchronized and are for the same duration, and if data sets are OPEN for the duration of the recording interval, information presented in Rule DAS400 will correspond well with device-related information (and application-related information) presented by other rules. If these conditions are not present, the information presented by Rule DAS400 might not correspond well with information presented elsewhere by the DASD Component.

Please send a note to [Don\\_Deese@cpexpert.com](mailto:Don_Deese@cpexpert.com) if these problems become troublesome. Perhaps the logic can be improved (by prorating the DASD42DS information, for example).

**Suggestion:** You should use the information displayed by Rule DAS400 to assess the data access characteristics for DFSMS-managed data sets residing on the device(s) with the most potential for performance improvement.